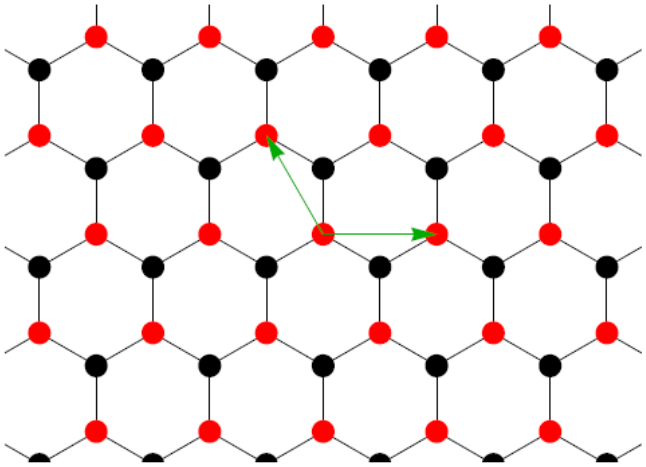
4. Topological phases-c

The Haldane Chern insulator

In the last section we saw how it is possible to obtain a quantum Hall state by first creating a Dirac cone, adding a mass term to it and finally to make this mass change sign. Following this recipe we were able to obtain chiral edge states without applying an external magnetic field. There is a real two-dimensional system which has Dirac cones: graphene. So in this section we will take graphene and make it into a topological system with chiral edge states. *A recipe to obtain a quantum Hall system was to first obtain a Dirac cone, add a mass term to it and finally to make this mass change sign. Following this recipe we were able to obtain chiral edge states without applying an external magnetic field.*

Graphene is a single layer of carbon atoms arranged in a honeycomb lattice. It is a triangular lattice with two atoms per unit cell, represented by the red and black atoms below. In this figure, we can see that there are two sites per unit cell (two sublattices) a and b (red and black respectively). The green arrow indicates the lattice vectors $\overrightarrow{v_1}$ and $\overrightarrow{v_2}$. If we shift the lattices by $m\overrightarrow{v_1} + n\overrightarrow{v_2}$ with m and n being integers, the lattice is invariant.



$$V(\vec{r}) = V(\vec{r} + m\overrightarrow{v_1} + n\overrightarrow{v_2})$$

Define a to the length of the nearest-neighbor (NN) bonds. Then we can show that $\overrightarrow{v_1} = (\sqrt{3}a, 0)$ and $\overrightarrow{v_2} = (-\sqrt{3}/2a, 3/2a)$. If we only consider hoppings between the NN sites, the Hamiltonian is

$$H = -t \sum_{<i,j>} a_i^\dagger b_j - t \sum_{<i,j>} b_i^\dagger a_j$$

where $< i, j >$ means nearest neighbors.

Band structures

There are three type of NN bonds: (1) along the y axis with $\theta = \pi/2$, (2) along $\theta = \pi/2 + 2\pi/3 = 7\pi/6$, (3) along $\theta = \pi/2 + 4\pi/3 = 11\pi/6$, and we need to write them out separately.

$$H = -t\sum_i a^\dagger_{\vec{r_i}} b_{r_i+\vec{e_1}} - t\sum_i a^\dagger_{\vec{r_i}} b_{r_i+\vec{e_2}} - t\sum_i a^\dagger_{\vec{r_i}} b_{r_i+\vec{e_3}} + h.c.$$

with $\vec{e_1} = (0, a)$, $\vec{e_2} = \left(-\sqrt{3}/2a, -a/2\right)$, $\vec{e_3} = \left(\sqrt{3}/2a, -a/2\right)$. Here the sum Σ_i sums over all unit cells (or say all red sites). If we go to the momentum space, by performing the 2D Fourier transformation,

$$a_k = \frac{1}{\sqrt{N}}\sum_i a_i\, e^{i\vec{k}\cdot\vec{r}}$$

$$a_i = \frac{1}{\sqrt{N}}\sum_k a_k\, e^{-i\vec{k}\cdot\vec{r}}$$

$$b_k = \frac{1}{\sqrt{N}}\sum_i a_i\, e^{i\vec{k}\cdot\vec{r}}$$

$$b_i = \frac{1}{\sqrt{N}}\sum_k a_k\, e^{-i\vec{k}\cdot\vec{r}}$$

the first term in the Hamiltonian becomes

$$\begin{aligned}
-t\sum_i a^\dagger_{\vec{r_i}} b_{r_i+\vec{e_1}} &= -\frac{t}{N}\sum_i\sum_k\sum_{k'} a^\dagger_k e^{i\vec{k}\cdot\vec{r_i}} b_{k'} e^{-i\vec{k'}\cdot(\vec{r_i}+\vec{e_1})} \\
&= -t\sum_k\sum_{k'} a^\dagger_k b_{k'} e^{-i\vec{k'}\cdot\vec{e_1}} \frac{1}{N}\sum_i e^{i\left(\vec{k}-\vec{k'}\right)\cdot\vec{r_i}} = -t\sum_k\sum_{k'} a^\dagger_k b_{k'} e^{-i\vec{k'}\cdot\vec{e_1}}\,\delta_{k,k'} \\
&= -t\sum_k a^\dagger_k b_k e^{-i\vec{k}\cdot\vec{e_1}}
\end{aligned}$$

Notice that, as we proved early on, we get a phase factor $\exp\left[i\vec{k}\cdot(\vec{r_a}+\vec{r_b})\right]$, which only depends on the separation between *a* and *b* sites. If we repeat the same procedure for the other two bonds, we get

$$\begin{aligned}
H = -t\sum_i a^\dagger_{\vec{r_i}} b_{r_i+\vec{e_1}} - t\sum_i a^\dagger_{\vec{r_i}} b_{r_i+\vec{e_2}} - t\sum_i a^\dagger_{\vec{r_i}} b_{r_i+\vec{e_3}} + h.c. \\
= -t\sum_k a^\dagger_k b_k\left(e^{-i\vec{k}\cdot\vec{e_1}} + e^{-i\vec{k}\cdot\vec{e_2}} + e^{-i\vec{k}\cdot\vec{e_3}}\right) \\
- t\sum_k b^\dagger_k a_k\left(e^{i\vec{k}\cdot\vec{e_1}} + e^{i\vec{k}\cdot\vec{e_2}} + e^{i\vec{k}\cdot\vec{e_3}}\right) \\
= \sum_k \begin{pmatrix} a^\dagger_k & b^\dagger_k \end{pmatrix} \begin{pmatrix} 0 & \mathcal{H}_{12}(\vec{k}) \\ \mathcal{H}_{21}(\vec{k}) & 0 \end{pmatrix} \begin{pmatrix} a_k \\ b_k \end{pmatrix}
\end{aligned}$$

$$\mathcal{H}_{12}(\vec{k}) = -t\left[e^{-i\vec{k}\cdot\vec{e_1}} + e^{-i\vec{k}\cdot\vec{e_2}} + e^{-i\vec{k}\cdot\vec{e_3}}\right]$$

$$\mathcal{H}_{21}(\vec{k}) = \mathcal{H}_{12}(\vec{k})^* = -t\left[e^{i\vec{k}\cdot\vec{e_1}} + e^{i\vec{k}\cdot\vec{e_2}} + e^{i\vec{k}\cdot\vec{e_3}}\right]$$

The kernel of the Hamiltonian:

$$\mathcal{H}(\vec{k}) = \begin{pmatrix} 0 & \mathcal{H}_{12}(\vec{k}) \\ \mathcal{H}_{21}(\vec{k}) & 0 \end{pmatrix}$$

The eigenvalues of $\mathcal{H}(\vec{k})$ gives the dispersion relation

$$\epsilon_{\pm}(\vec{k}) = \pm|\mathcal{H}_{12}(\vec{k})| = \pm|t|\sqrt{3 + 2\cos(\sqrt{3}k_x a) + 4\cos\left(\frac{\sqrt{3}}{2}k_x a\right)\cos\left(\frac{3}{2}k_y a\right)}$$

The dispersion $\epsilon_{\pm}$ is a periodic function of $\vec{k}$ space (the hexagon repeats itself in the figure shown above). The first BZ is a hexagon. However, due to the periodic structure in k-space, the three corner points with $\theta = 0, 2\pi/3$ and $4\pi/3$ are the same point (their momenta differ by a reciprocal vector). Similarly, the three corners with $\theta = \pi, \pi + 2\pi/3$ and $\pi + 4\pi/3$ are the same point. Therefore, there are only two different corner points and they are known as the K and K' points, where

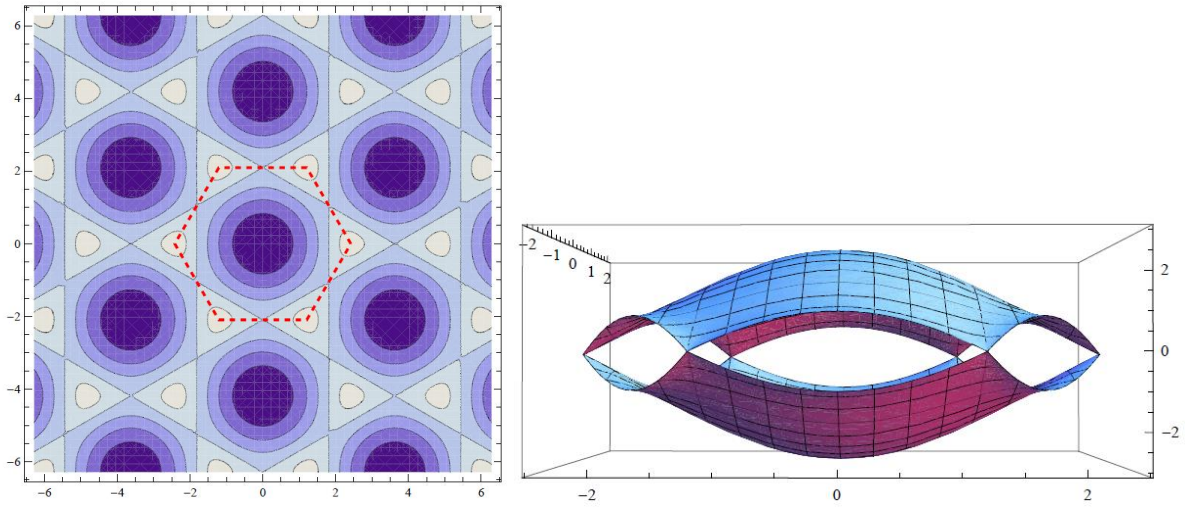
$$K = \left(\frac{4\pi}{3\sqrt{3}a}, 0\right), \quad K' = \left(-\frac{4\pi}{3\sqrt{3}a}, 0\right)$$

For most of the momentum points, one of the two bands have positive energy $\epsilon_+ > 0$ and the other one has negative energy $\epsilon_- < 0$. However, at the corner of the BZ, K and K', the two bands are degenerate $\epsilon_+ = \epsilon_- = 0$.

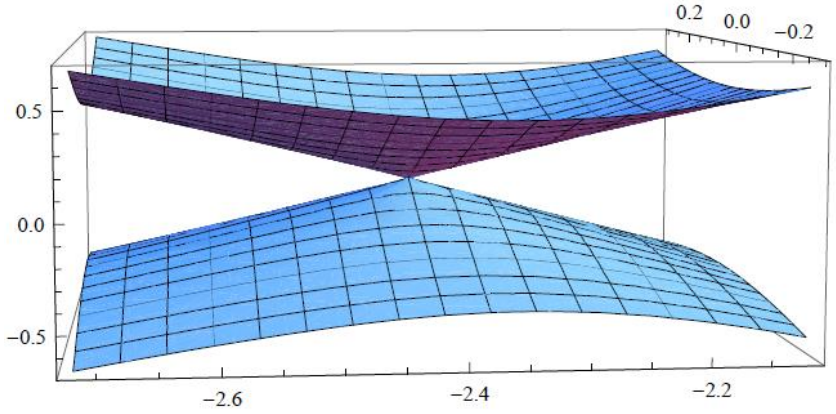
$$\begin{aligned}
\epsilon_{\pm}(\vec{K}) = \pm|\mathcal{H}_{12}(\vec{K})| &= \pm|t|\sqrt{3 + 2\cos(\sqrt{3}k_x a) + 4\cos\left(\frac{\sqrt{3}}{2}k_x a\right)\cos\left(\frac{3}{2}k_y a\right)} \\
&= \pm|t|\sqrt{3 + 2\cos\left(\sqrt{3}\frac{4\pi}{3\sqrt{3}a}a\right) + 4\cos\left(\frac{\sqrt{3}}{2}\frac{4\pi}{3\sqrt{3}a}a\right)\cos\left(\frac{3}{2}\times 0\right)} \\
&= \pm|t|\sqrt{3 + 2\cos\left(\frac{4\pi}{3}\right) + 4\cos\left(\frac{2\pi}{3}\right)} = \pm|t|\sqrt{3 + 2\times\left(-\frac{1}{2}\right) + 4\times\left(-\frac{1}{2}\right)} = 0
\end{aligned}$$

Two bands have the same energy at K and K', and they are in fact band crossing points.

Near K or K' points, the energies of the two bands are linear functions of momentum $\epsilon_{\pm} \propto (k - K)$. To see this, we expand $\epsilon_{\pm}(\vec{k})$ near $\vec{k}\sim\vec{K}$ and $\vec{k}\sim\vec{K'}$



$$\epsilon_{\pm}(\vec{k}) = \epsilon_{\pm}(\vec{K}+\vec{q})$$
$$= \pm|t|\sqrt{3 + 2\cos\left[\sqrt{3}\left(\frac{4\pi}{3\sqrt{3}a} + q_x\right)a\right] + 4\cos\left[\frac{\sqrt{3}}{2}\left(\frac{4\pi}{3\sqrt{3}a} + q_x\right)a\right]\cos\left(\frac{3}{2}q_y a\right)}$$
$$= \pm\frac{3}{2}|t|a\sqrt{q_x^2 + q_y^2} + O(q^2) = \pm\frac{3}{2}|t|aq + O(q^2)$$



Near the K point, we have

$$\mathcal{H}(k) = \mathcal{H}(K+q) = \frac{3}{2}ta\begin{pmatrix} 0 & q_x - iq_y \\ q_x + iq_y & 0 \end{pmatrix} + O(q^2) \approx \frac{3}{2}ta\left(q_x\sigma_x + q_y\sigma_y\right) = c\vec{q}\cdot\vec{\sigma}$$

Near the K' point, we have

$$\mathcal{H}(k) = \mathcal{H}(K' + q) = -\frac{3}{2}ta\begin{pmatrix} 0 & q_x + iq_y \\ q_x - iq_y & 0 \end{pmatrix} + O(q^2) \approx -\frac{3}{2}ta\left(q_x\sigma_x - q_y\sigma_y\right)$$
$$= -c\vec{q}\cdot(\sigma_x\vec{\sigma}\sigma_x)$$

Each of them is a Weyl fermion. These two Weyl fermions with opposite chirality forms a Dirac fermion. Dirac fermions we learned in quantum mechanics:

$$i\partial_t\psi = \begin{pmatrix} 0 & c\vec{q}\cdot\vec{\sigma} \\ c\vec{q}\cdot\vec{\sigma} & 0 \end{pmatrix}\psi$$

If one makes a unitary transformation (changing the basis, which doesn't change any physics)

$$\psi' = \begin{pmatrix} 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{2}} & 0 & 0 & \frac{1}{\sqrt{2}} \\ 0 & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ -\frac{1}{\sqrt{2}} & 0 & 0 & \frac{1}{\sqrt{2}} \end{pmatrix}\psi$$

This equation turns into

$$i\partial_t\psi' = \begin{pmatrix} -c\vec{q}\cdot(\sigma_x\vec{\sigma}\sigma_x) & 0 \\ 0 & c\vec{q}\cdot\vec{\sigma} \end{pmatrix}\psi'$$

The Dirac point (band crossing points) in a honeycomb lattice is stable as long as the space inversion symmetry $r \to -r$ and the time reversal symmetry $t \to -t$ are preserved. No matter how one perturb the systems (e.g. adding longer-range hoppings), the Dirac point is always there as long as the two symmetries mentioned above are preserved. For graphene, the lower band is filled and the upper band is empty, which is known as "half-filling". The "half" here means that the number of electrons N_e over the number of sites N_S is $N_e/N_S = 1/2$. However, it is worthwhile to notice that the ratio between N_e and the number of unit cells N is 1, because there are two sites in each unit cell. So one of the two bands are totally filled. By gating, one can tune the Fermi energy slightly away from the Dirac point, which is known as a dopped graphene.

Discrete symmetries of graphene

What are the symmetries of this Hamiltonian? First, sublattice symmetry. Graphene is the prototype of a system with sublattice symmetry, which makes the Hamiltonian block off-diagonal with respect to the two sublattices. The sublattice symmetry reads

$$\sigma_z\mathcal{H}(\vec{\mathrm{k}})\sigma_z = -\mathcal{H}(\vec{\mathrm{k}})$$

Sublattice symmetry is only approximate, and it is consequence of the nearest neighbor tight-binding model. It protects the Dirac points and needs to be broken in order to open a gap.

In addition to sublattice and inversion symmetry, the honeycomb lattice also has a three-fold rotation symmetry around the center of the unit cell. This symmetry is important to make the Dirac cones appear in the first place, but it will not play a role in all that follows.

Finally, there is time-reversal symmetry. Since we are not considering the spin degree of freedom of the electrons, the time-reversal symmetry operator in real space is just complex conjugation. In momentum space representation, time-reversal symmetry reads

$$\mathcal{H}(\vec{\mathrm{k}}) = \mathcal{H}^*(-\vec{\mathrm{k}})$$

It's important to note that time-reversal symmetry sends K into K′ and therefore it exchanges the two Dirac cones.

The product of (approximate) sublattice and time-reversal symmetries yields a further discrete symmetry, a particle-hole symmetry

$$\sigma_z \mathcal{H}^*(-\vec{\mathrm{k}})\sigma_z = -\mathcal{H}(\vec{\mathrm{k}})$$

Making graphene topological

Let's recall that our goal is to make our graphene sheet enter a quantum Hall state, with chiral edge states. The first necessary step is to make the bulk of the system gapped. How can we open a gap in graphene? The Dirac points are protected by both sublattice (inversion) and time-reversal symmetry. So there are many ways we can think of to open an energy gap at K and K'.

First try: The easiest way to break sublattice symmetry is to assign an opposite onsite energy M or $-M$ to the *A* or *B* sites respectively. The Hamiltonian is then given by

$$\mathcal{H}(\vec{k}) + M\sigma_z$$

This leads to a gapped spectrum,

$$\epsilon_{\pm}(\vec{k}) = \pm\sqrt{\left|\mathcal{H}_{12}(\vec{k})\right|^2 + M^2}$$

However, we quickly realize that by doing this we end up in a rather boring situation. Taking the limit $|M| \gg t$, we obtain electronic states which are localized in one of the two sublattices *a* or *b*, independent of the sign of M. Most importantly, there is no trace of edge states.

It's easy to see why this mass term is hopeless: it preserves time-reversal symmetry. And with the time-reversal symmetry present, it is definitely impossible to obtain chiral edge states.

Aharonov–Bohm effect and complex hopping

Can hopping strength t be complex? The answer is Yes, and the phase can come from the Aharonov–Bohm effect in the presence of a magnetic field or spin-orbital couplings.

Now, let's consider a discrete lattice. Consider three sites *a*, *b* and *c*. The hopping strength between these three sites are tab t_{ab}, t_{bc} and t_{ca} respectively. If a particle hops from *a* to *b* and then to *c*, the hopping strength around this loop is:

$$t_{ab}t_{bc}t_{ca} = |t_{ab}|e^{i\phi_{ab}} * |t_{bc}|e^{i\phi_{bc}} * |t_{ca}|e^{i\phi_{ca}} = |t_{ab}t_{bc}t_{ca}|e^{i(\phi_{ab}+\phi_{bc}+\phi_{ca})}$$

The phase picked up by the electron is:

$$\phi_{ab} + \phi_{bc} + \phi_{ca} = e/\hbar \iint B \cdot dS$$

If B is nonzero inside the triangle formed by these three sites, the phase for these hoppings are nonzero.

Please notice that (1) t_{ab} and t_{ba} has opposite phase, due to the Hermitian condition $t_{ab} = t_{ba}^*$

(2) The individual phases for t_{ab}, t_{bc} and t_{ca} have no physical meaning and their phases depend on the gauge choice. However, the total phase around a loop is independent of gauge, and in fact it is a physical observable, i.e. the magnetic flux.

Proof:

We know that the phase of the hopping term is

$$\phi_{ab} = e/\hbar \int_a^b A \cdot dl$$

Underage transformation: $A \to A + \nabla\chi$

$$\phi_{ab} \to \phi'_{ab} = e/\hbar \int_a^b (A + \nabla\chi) \cdot dl = e/\hbar \int_a^b A \cdot dl + e/\hbar \int_a^b \nabla\chi \cdot dl = \phi_{ab} + (\chi_b - \chi_a)e/\hbar$$

Obviously, ϕ_{ab} is not a physical observable, since it relies on the gauge choice. However, the total phase around a loop is different. It is a loop integral of A, which is gauge independent and the physics meaning of this Integral is the magnetic flux.

$$\phi_{ab} + \phi_{bc} + \phi_{ca} = e/\hbar \oint A \cdot dl$$

$$\to \phi'_{ab} + \phi'_{bc} + \phi'_{ca} = e/\hbar \oint (A + \nabla\chi) \cdot dl = e/\hbar \oint A \cdot dl + e/\hbar \oint \nabla\chi \cdot dl = e/\hbar \oint A \cdot dl$$
$$= \phi_{ab} + \phi_{bc} + \phi_{ca}$$

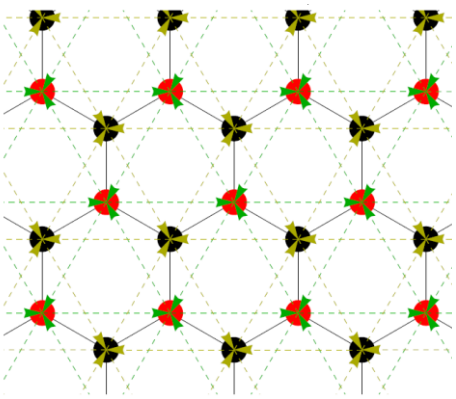
The complex hopping strength induced by B fields breaks the time-reversal symmetry because $B \to -B$ under time reversal. (in other words, under time-reversal one needs to flip the sign of all these phases).

Second try: There is another, more ingenious way to gap out the Dirac cones in graphene, which is the essence of today's model. It involves **adding imaginary next-nearest neighbor hoppings**.

In his seminal paper, Haldane considered a honeycomb model with no net magnetic flux but with complex phases $e^{\pm i\phi}$ on the next-to-nearest neighbor hoppings. But the field is periodic and have zero net flux per hexagon. Haldane model describes the model of graphene with **real nearest-neighbor-hopping parameters** but **complex next-nearest-neighbor-hopping parameters** resulting from a nonzero magnetic field.

Complex next-nearest-neighbor (NNN) hoppings (breaking T-symmetry using B fields)

Now let us add some NNN hoppings and assume their hopping amplitudes are complex. For simplicity, we choose the amplitude and the complex phase to be the same for all NNN bonds. If the hopping is along the arrows marked in the figure, the phase of the hopping strength is ϕ. If one hops in the opposite direction, the phase is $-\phi$.



With the direction of the arrow, we denote the direction in which the hopping is $+it'$(it is $-it'$ in the opposite direction). Note the following things about these hoppings:

- They are purely imaginary and, furthermore, they all have the same chirality, in the sense that they all follow the orientation of your right hand, if the thumb points out from the screen.
- They couple sites of same type: *a* with *a* and *b* with *b*.

These characteristics tell us that the new hoppings break both time-reversal symmetry and sublattice symmetry. In this model, both inversion symmetry and time-reversal symmetry are simultaneously broken in graphene. **Inversion symmetry is broken by assigning different on-site energies to the two inequivalent sublattices of the honeycomb lattice (to be shown later), while time-reversal invariance is lifted by local magnetic fluxes organized so that the net flux per unit cell vanishes**. Therefore, the NN hopping amplitudes are not affected by the magnetic fluxes, whereas NNN hopping amplitudes acquire an Aharonov–Bohm phase.

This complex phase can be realized (in theory) by applying a staggered *B* field, which is positive near the center of each hexagon and negative near the edges. The NNN hoppings are from an *a*-site to another *a*-site (and from a *b*-site to another *b*-site). For *a*-to-*a* hoppings, there are three different types of NNN bonds, along $\theta = 0, 2\pi/3$ and $4\pi/3$. Same is true for *b*-to-*b* hoppings.

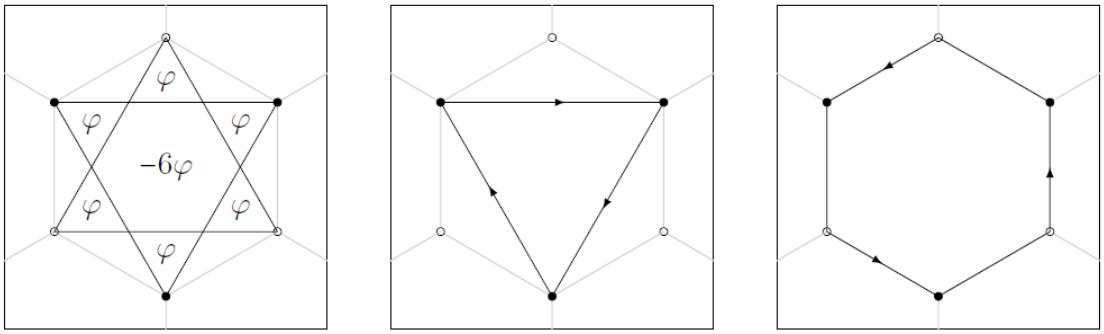
Moreover, the Aharanov–Bohm phases due to the time-reversal breaking local magnetic fluxes are taken into account through the Peierls substitution:

$$t_{ij} \rightarrow t_{ij} \exp\left(-i\frac{e}{\hbar}\int_{\Gamma_{ij}} \mathbf{A}\cdot d\boldsymbol{l}\right)$$

where t_{ij} is the hopping parameter between sites i and j, and where Γ_{ij} is the hop trajectory from site i to j and $\boldsymbol{A}$ is a potential vector accounting for the presence of the magnetic flux. In Haldane's model, magnetic fluxes are imposed such that the phase accumulated through a NN $a \rightarrow b$ (or $b \rightarrow a$) hopping vanishes, whereas the phase accumulated through a NNN hopping $a \rightarrow a$ or $b \rightarrow b$ is nonzero (see the figure below for a possible flux distribution). Notice that to use Peierls substitution, we choose a gauge for the vector potential; he Peierls substitution amounts to the substitution:

$$t \rightarrow t \quad \text{and} \quad t' \rightarrow t' e^{i\phi}$$

where the Aharonov–Bohm phase - due to the local magnetic flux is taken as a parameter of the model.



Example of a choice for magnetic flux in the Haldane cell (left). We have used $\varphi = \phi/2$ to simplify. NNN corresponds to a nonzero flux (middle), whereas the NN gives a zero flux (right), so the total flux through the hexagon is zero.

So the Hamiltonian is:

$$H_{NNN} = -t' e^{i\phi} \sum_i a^\dagger_{r_i} a_{r_i+v_1} - t' e^{i\phi} \sum_i a^\dagger_{r_i+v_1} a_{r_i-v_3} - t' e^{i\phi} \sum_i a^\dagger_{r_i-v_3} a_{r_i} + h.c. + (a \rightarrow b \ \ and \ \phi \rightarrow -\phi)$$

Here, $\overrightarrow{v_1} = (\sqrt{3}a, 0)$ and $\overrightarrow{v_2} = (-\sqrt{3}/2a, 3/2a)$ as shown before.

$$\begin{aligned} H_{NNN} &= -t' e^{i\phi} \sum_k a^\dagger_k a_k \left(e^{-i\vec{k}\cdot\overrightarrow{v_1}} + e^{-i\vec{k}\cdot\overrightarrow{v_2}} + e^{-i\vec{k}\cdot\overrightarrow{v_3}}\right) + h.c. + (a \rightarrow b \ \ and \ \phi \rightarrow -\phi) \\ &= -2t' \sum_k a^\dagger_k a_k [\cos(k\cdot v_1 - \phi) + \cos(k\cdot v_2 - \phi) + \cos(k\cdot v_3 - \phi)] \\ &\quad - 2t' \sum_k b^\dagger_k b_k [\cos(k\cdot v_1 - \phi) + \cos(k\cdot v_2 - \phi) + \cos(k\cdot v_3 - \phi)] \end{aligned}$$

$$H = H_{NN} + H_{NNN} = \sum_k \begin{pmatrix} a_k^\dagger & b_k^\dagger \end{pmatrix} \begin{pmatrix} \mathcal{H}_{11}(k) & \mathcal{H}_{12}(k) \\ \mathcal{H}_{21}(k) & \mathcal{H}_{22}(k) \end{pmatrix} \begin{pmatrix} a_k \\ b_k \end{pmatrix}$$

$$\mathcal{H}_{12}(k) = -t\left(e^{-i\vec{k'}\cdot\vec{e_1}} + e^{-i\vec{k'}\cdot\vec{e_2}} + e^{-i\vec{k'}\cdot\vec{e_3}}\right)$$

$$\mathcal{H}_{21}(k) = \mathcal{H}_{12}(k)^* = -t\left(e^{i\vec{k'}\cdot\vec{e_1}} + e^{i\vec{k'}\cdot\vec{e_2}} + e^{i\vec{k'}\cdot\vec{e_3}}\right)$$

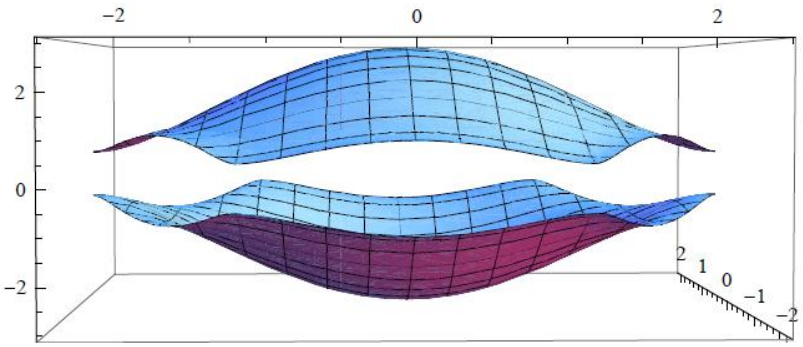
$$\mathcal{H}_{11}(k) = -2t'[\cos(k\cdot v_1 - \phi) + \cos(k\cdot v_2 - \phi) + \cos(k\cdot v_3 - \phi)]$$

$$\mathcal{H}_{22}(k) = -2t'[\cos(k\cdot v_1 + \phi) + \cos(k\cdot v_2 + \phi) + \cos(k\cdot v_3 + \phi)]$$

If one computes the eigenvalues of $\mathcal{H}(k)$, one finds that the two bands never cross with each other for any k (if t' is non-zero and ϕ is NOT an integer multiplied by π).

Using Pauli matrices:

$$\mathcal{H} = \mathcal{H}_0(k)I + \mathcal{H}_x(k)\sigma_x + \mathcal{H}_y(k)\sigma_y + \mathcal{H}_z(k)\sigma_z$$



The energy dispersion for both bands ($\epsilon_\pm$ as a function of kx and ky) in the first BZ (within the red dash lines mark in the figure above).

$$\mathcal{H}_0 = \frac{1}{2}\left(\mathcal{H}_{11}(k) + \mathcal{H}_{22}(k)\right) = -2t'\cos\phi\,[\cos(k\cdot v_1) + \cos(k\cdot v_2) + \cos(k\cdot v_3)]$$

$$\mathcal{H}_z = \frac{1}{2}\left(\mathcal{H}_{11}(k) - \mathcal{H}_{22}(k)\right) = -2t'\sin\phi\,[\sin(k\cdot v_1) + \sin(k\cdot v_2) + \sin(k\cdot v_3)]$$

$$\mathcal{H}_x = Re[\mathcal{H}_{21}(k)] = -t\left[\cos\left(\vec{k}\cdot\vec{e_1}\right) + \cos\left(\vec{k}\cdot\vec{e_2}\right) + \cos\left(\vec{k}\cdot\vec{e_3}\right)\right]$$

$$\mathcal{H}_y = Im[\mathcal{H}_{21}(k)] = -t\left[\sin\left(\vec{k}\cdot\vec{e_1}\right) + \sin\left(\vec{k}\cdot\vec{e_2}\right) + \sin\left(\vec{k}\cdot\vec{e_3}\right)\right]$$

The energy dispersion:

$$\epsilon_\pm(\vec{k}) = \mathcal{H}_0(\vec{k}) \pm \sqrt{\mathcal{H}_x(\vec{k})^2 + \mathcal{H}_y(\vec{k})^2 + \mathcal{H}_z(\vec{k})^2}$$

The gap between the two bands:

$$\epsilon_+(\vec{k}) - \epsilon_-(\vec{k}) = 2\sqrt{\mathcal{H}_x(\vec{k})^2 + \mathcal{H}_y(\vec{k})^2 + \mathcal{H}_z(\vec{k})^2}$$

At K or K', $\mathcal{H}_x(\vec{k}) = \mathcal{H}_y(\vec{k}) = 0$ and the gap is:

$$\epsilon_+(\vec{k}) - \epsilon_-(\vec{k}) = 2\left|\mathcal{H}_z(\vec{k} = K)\right| = 6\sqrt{3}t' \sin\phi$$

In fact, at the K point,

$$\mathcal{H}_z = -3\sqrt{3}t' \sin\phi$$

At the K' point

$$\mathcal{H}_z = 3\sqrt{3}t' \sin\phi$$

They have opposite signs (as long as ϕ is not $n\pi$). Based on what we learned last time, this means that one cannot define the wavefunction in the whole BZ. We need to cut the systems into two regions. The region I contains the K point, the region II contains the K' points. And we need to use different eigenvectors for these two regions. Using the same method we learned in the last class, one finds that the Chern number here is ±1.

Here, we can draw a small circle around the K point. Inside the circle, we use wavefunction $u^{(I)}$. Outside it, we use $u^{(II)}$.

$$e^{i\phi(k)} = \frac{\dfrac{\mathcal{H}_z(k) + \left|\vec{\mathcal{H}}(k)\right|}{\mathcal{H}_x(k) + i\mathcal{H}_y(k)}}{\left|\dfrac{\mathcal{H}_z(k) + \left|\vec{\mathcal{H}}(k)\right|}{\mathcal{H}_x(k) + i\mathcal{H}_y(k)}\right|} = \frac{\left|\mathcal{H}_x(k) + i\mathcal{H}_y(k)\right|}{\mathcal{H}_x(k) + i\mathcal{H}_y(k)} = \frac{\left|q_x + iq_y\right|}{q_x + iq_y} = \frac{1}{e^{i\theta}} = e^{-i\theta}$$

Potential energy (breaking inversion symmetry)

As we said before, i**nversion symmetry is broken by assigning different on-site energies to the two inequivalent sublattices of the honeycomb lattice.** Let's keep NNN hoppings to be zero for now and add some potential energy to the Hamiltonian.

$$H_{Potential} = (V + M)\sum_i a_i^\dagger a_i + (V - M)\sum_i b_i^\dagger b_i = VN + M\sum_i a_i^\dagger a_i - M\sum_i b_i^\dagger b_i$$

The V part (average potential between a and b sites) just adds a constant term to the energy, since the total particle number N is conserved. So we can drop the V term and only consider the difference between the potential energies at a and b sites (M). The sublattice symmetry breaking term with on-site energies $+M$ for sites of sublattice a, and $-M$ for sublattice b, which thus breaks inversion symmetry of the honeycomb sublattices.

In k-space,

$$H_{Potential} = M\sum_i a_i^\dagger a_i - M\sum_i b_i^\dagger b_i = M\sum_k a_k^\dagger a_k - M\sum_k b_k^\dagger b_k = \sum_k \begin{pmatrix} a_k^\dagger & b_k^\dagger \end{pmatrix} M\sigma_z \begin{pmatrix} a_k \\ b_k \end{pmatrix}$$

It adds a σ_z component to the Hamiltonian. Same as the NNN complex hopping, this term also opens a gap at the Dirac points. At K and K', $\mathcal{H}$ now has the same sign ($\mathcal{H} = M$ for any k). So, we can use one wave-function for the whole BZ and thus the system is a trivial insulator.

The interplay between the TR-breaking mass term and the inversion-breaking mass gives the Haldane Hamiltonian, a nontrivial structure with three phases of Hall conductance, 0, 1, −1, very much like in our square-lattice example. Let us see how. We could directly compute the Hall conductance of the system using the Jacobian formula, but that would not be very physically illuminating. Instead, let us again use the Dirac argument.

What will happen if we have both $H_{Potential}$ and H_{NNN}?

We just need to look at the signs of $\mathcal{H}_z$ at K and K' points. If they have the same sign, we can use one wavefunction to cover the whole BZ, so C = 0 (trivial insulator). If they have opposite signs, the system is a topological insulator with C=±1.

At the K point,

$$\mathcal{H}_z = M - 3\sqrt{3}t' \sin\phi$$

At K' point,

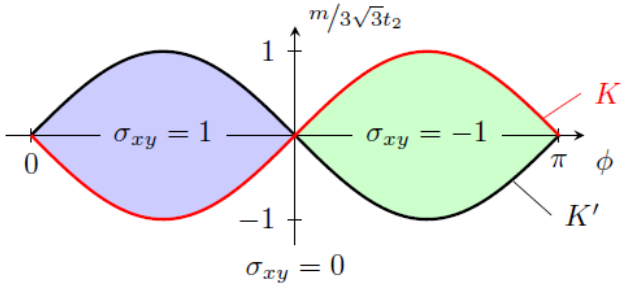
$$\mathcal{H}_z = M + 3\sqrt{3}t' \sin\phi$$

Therefore, as long as $|M| < \left|3\sqrt{3}t' \sin\phi\right|$, the system is a topological insulator ($\mathcal{H}_z$ flips sign). If $|M| > \left|3\sqrt{3}t' \sin\phi\right|$, $\mathcal{H}_z$ is always positive (or negative) and thus the system is topologically trivial. The marginal case $|M| = \left|3\sqrt{3}t' \sin\phi\right|$ is a topological transition, where $\mathcal{H}_z = 0$ at either the K point or the K' point. Because $\mathcal{H}_x = \mathcal{H}_y = 0$ at these two points, the gap must be zero at one of the two points.

Phase Diagram of the Haldane Hamiltonian

Note that the gap-closings at K and K' happen at different values of M (for $\phi \neq 0, \pi$). To obtain the value of the Hall conductance, as we did before, we have to start from a phase where we know the its value and trace the changes in the Chern number as we undergo gap closing and reopening transitions. Such a phase is easy to find: we start from the case $M \to \pm\infty$. In that case the wavefunction is localized on the site a or b and the system is trivial, with eigenstates constant in momentum space. As such, the system has no Hall conductance. Moreover, the two Dirac points give rise to a change in σ_{xy} of opposite sign has Det(A) as a different sign.

To determine the phase diagram, let us find the points in the parameter space where the local gap closes (i.e. $h = 0$) at some points of the Brillouin torus. In graphene, which corresponds to $(M, \phi) = (0, 0)$ in the diagram, the two energy bands are degenerate ($h = 0$) at the Dirac points. At a generic point of the diagram, this degeneracy is lifted, and the system is an insulator ($h \neq 0$) except when $|M| = 3\sqrt{3}t'\,sin(\phi)$. The corresponding line separates four a priori different insulating states.



The Haldane Hamiltonian is similar to the Hamiltonian of $h(\boldsymbol{k}) = \sum_{a,b=1}^{2} k_a A_{ab} \sigma_b + M\sigma_3$ and thus its Hall conductance can be readily obtained by using the Determinant Formula for the Hall Conductance of a generic Dirac Hamiltonian $\sigma_{xy} = \frac{1}{2}\frac{e^2}{h} sign(M) sign(det(A))$. Without loss of generality, we now assume $\phi > 0$ with conditions of $-\pi < \phi < \pi$, $t' > 0$ and start lowering m from $+\infty$. At $m = 3\sqrt{3}t'\,sin(\phi)$, the K Dirac fermion goes through a gap-closing-and-opening transition and the Hall conductance changes from $\frac{1}{2}\frac{e^2}{h}\text{sign}\big(M - 3\sqrt{3}t'\sin(\phi)\big) = \frac{1}{2}\frac{e^2}{h}$ to $-\frac{1}{2}\frac{e^2}{h}$, with a change of Hall conductance equal to -1. The gap at K' stays open for $M = 3\sqrt{3}t'\sin(\phi)$, and only K goes through gap-closing-and-reopening transition to give a phase with

$$\sigma_{xy} = -1$$

As we lower m even more, we reach the value where

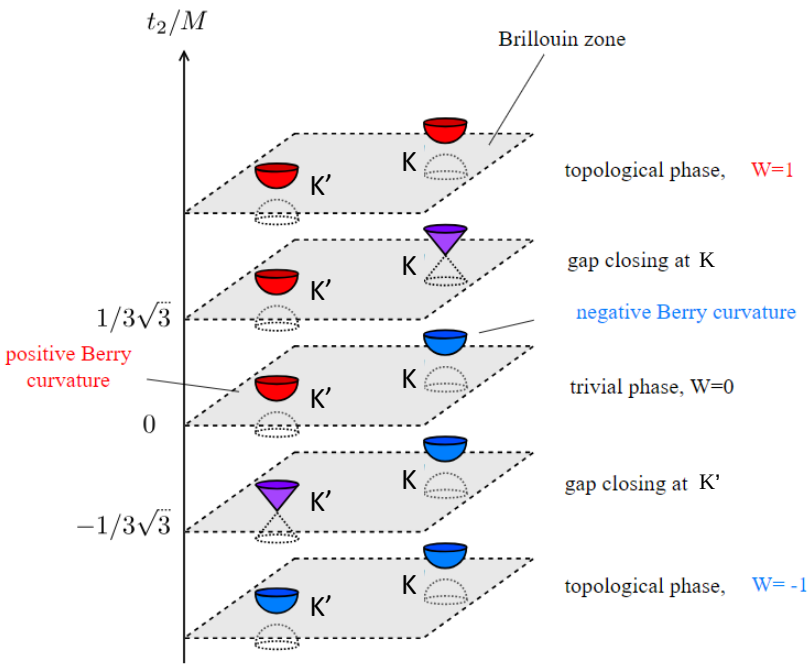
$$M = -3\sqrt{3}\mathrm{t}'\sin(\phi)$$

at which point the K' Dirac fermion goes through a gap-closing and re-opening transition, which changes the Hall conductance by $\Delta\sigma_{xy} = 1\frac{e^2}{h}$ (the point K remains gapped). The change in Hall conductance means that the phase $\sigma_{xy} = -1\frac{e^2}{h}$ before the gap closing goes to a phase with $\sigma_{xy} = 0$ after the gap closing and reopening. This is also necessary because the phase $M < -3\sqrt{3}t'\sin(\phi)$ is adiabatically continuable to the state $M \to \infty$, which must have zero Hall conductance. In short, we have

- $M > 3\sqrt{3}\mathrm{t}'\sin(\phi)$: $\sigma_{xy} = 0$
- $-3\sqrt{3}\mathrm{t}'\sin(\phi) < M < 3\sqrt{3}\mathrm{t}'\sin(\phi)$ for $\phi > 0$: At $M = 3\sqrt{3}\mathrm{t}'\sin(\phi)$ the gap closes at K and we have a $\Delta\sigma_{xy} = -\frac{e^2}{h}$. The gap at K' stays open.
- $M \leq -3\sqrt{3}\mathrm{t}'\sin(\phi)$ for $\varphi > 0$: The gap at K' closes and hence the Chern number changes back to 0.

For $\phi < 0$ the signs of the Chern numbers are inverted.

These points evolve continuously with the parameters (M, ϕ). From this perspective, the pure graphene $(M, \phi) = (0, 0)$ corresponds to a bicritical transition where the gap closes simultaneously at both points K, K' which are the two nonequivalent Dirac points.



Let draw a full transition diagram. First we need to know how does time-reversal symmetry influence the Berry curvature? The Berry curvature and momentum change sign under time-reversal, so that the Berry curvature at one momentum becomes opposite to the Berry curvature at opposite momentum. As shown in the figure on the right, the two massless Dirac cones appearing for $t' = \pm M/3\sqrt{3}$ are the sources of the Berry curvature, which then "spreads" along the vertical axis, passing through the Brillouin zones of the gapped phases.

Remarks:

- It demonstrates that topological insulator is a generic concept, which may appear in any insulating systems (NOT just quantum Hall).
- It also demonstrates that as long as the topological index is nonzero, one will observe all the topological phenomena expected for a quantum Hall state, including the quantized Hall conductivity and the existence of the edge states.
- The key differences between the model of Haldane and the quantum Hall effects are (1) the B field is on average zero in the model of Haldane while the QHE has a uniform B field and (2) there is a very strong lattice background in the model of Haldane while the QHE requires weak lattice potential.
- The model of Haldane is also the foundation to explore more complicated and exotic topological states. For example, the time-reversal invariant topological insulator was first proposed using a modified Haldane's models, which we will study latter in the semester.